ELSEVIER

Gene 255 (2000) 245–255

GENE
AN INTERNATIONAL JOURNAL ON GENES, GENOMES AND EVOLUTION

www.elsevier.com/locate/gene

Isolation and characterization of disease resistance gene homologues from rice cultivar IR64

Lawrence L. Ilag [a,b,1], Ram C. Yadav [c,2], Ning Huang [b,3], Pamela C. Ronald [c], Frederick M. Ausubel [a,*]

[a] *Department of Genetics, Harvard Medical School and Department of Molecular Biology, Massachusetts General Hospital, Boston, MA 02114, USA*
[b] *Genome Mapping Laboratory, Department of Plant Breeding, Genetics, and Biochemistry, International Rice Research Institute, P.O. 933, 1099 Manila, Philippines*
[c] *Department of Plant Pathology, University of California, Davis, CA 95616, USA*

Received 26 April 2000; received in revised form 19 June 2000; accepted 17 July 2000
Received by D. Baulcombe

Abstract

We initiated a search for disease resistance (*R*) gene homologues in rice cultivar IR64, one of the most agronomically important rice varieties in the world, with the assumption that some of these homologues would correspond to previously identified disease resistance loci. A family of rice *R* gene homologues was identified using the *Arabidopsis* NBS–LRR disease resistance gene *RPS2* as a hybridization probe. Because member genes of this rice *R* gene family exhibit features characteristic of the NBS–LRR class of resistance genes, the family was given the name *NRH* (for NBS–LRR resistance gene homologues). Three members of the *NRH* family, *NRH1*, *NRH2*, and *NRH3*, were cloned and studied in detail. In IR64, *NRH1* and *NRH2* appear to encode full-length polypeptides, whereas *NRH3* is prematurely truncated with a stop codon generated by a frameshift. *NRH1* maps on chromosome 5, and *NRH2* and *NRH3* are less than 48 kb apart on chromosome 11. Although *NRH1*, *NRH2*, and *NRH3* map to regions of the rice genome where disease resistance loci to *Xanthomonas oryzae* pv. *oryzae* (*Xoo*) have been identified, susceptible rice varieties transformed with either *NRH1* or *NRH2* failed to exhibit increased resistance to a set of well-characterized *Xoo* strains.

Keywords: NBS–LRR; *NRH*; *Xa4*; *Xanthomonas*

1. Introduction

A major advance towards understanding the molecular events that result in plant disease resistance has been the cloning of resistance (*R*) genes from a variety of plant species that confer resistance to a variety of fungal, bacterial, and viral pathogens. It is remarkable that, despite their specificity, almost all *R* genes identified to date encode polypeptides that share similar structural motifs, allowing them to be classified into five broad categories.

The maize resistance gene *HM1*, which represents the first class, encodes a reductase that detoxifies the HC toxin, enabling the fungus *Cochliobolus carbonum* race 1 to colonize *Hm1*-deficient maize cultivars (Johal and Briggs, 1992). The other four classes of *R* genes encode proteins that are generally thought to serve as receptors that directly or indirectly recognize signals generated by corresponding pathogen avirulence (*avr*) genes. The second class includes the tomato *Pto* gene, which encodes a protein that has a serine–threonine protein kinase domain (Martin et al., 1993). The third and largest class is the NBS–LRR class, which includes the *Arabidopsis RPS2* and *RPM1* genes, and encodes proteins that have in common N-terminal nucleotide binding sites (NBS) and C-terminal leucine-rich repeats

Abbreviations: GUS, β-glucuronidase; NBS–LRR, nucleotide binding site–leucine-rich repeat; QTL, quantitative trait locus; R, resistance; *Xoo*, *Xanthomonas oryzae* pv. *oryzae*.

* Corresponding author. Tel.: +1-617-726-5969; fax: +1-617-726-5949.
E-mail address: ausubel@frodo.mgh.harvard.edu (F.M. Ausubel)
[1] Present address: Dupont Agricultural Products, Lancaster Pike and Rt. 141, Barley Mill Plaza, Bldg. 16, Wilmington, DE 19805, USA.
[2] Present address: Department of Biotechnology and Molecular Biology, CCS Haryana Agril University, Hisar 125004, India.
[3] Present address: Applied Phytologics, Inc., 4110 N. Freeway Blvd., Sacramento, CA 95834, USA.

0378-1119/00/$ - see front matter
PII: S0378-1119(00)00333-4

(LRR) that may be involved in protein–protein interactions. The fourth class includes the tomato *Cf* genes that encode putative transmembrane receptors with LRRs making up most of the extracellular domain (Jones et al., 1994). The fifth class is represented by the rice *Xa21* gene that encodes a putative receptor kinase, having both extracytoplasmic LRR domains similar to those of the *Cf*-encoded proteins, and an intracellular kinase domain like that of Pto (Song et al., 1995). Genes belonging to the latter four classes appear to play a role in signal transduction, suggesting that pathogen recognition and disease resistance occur through multicomponent signaling pathways.

In rice, a number of disease resistance genes have been cloned, including *Xa21*, *Xa1*, and *Pib*. *Xa21* and *Xa1* confer resistance to *Xanthomonas oryzae* pv. *oryzae* (*Xoo*), causal agent of bacterial leaf blight, the most destructive bacterial disease of rice in Asia and Africa, while *Pib* confers resistance to the fungal pathogen *Magnaporthe grisea*, causal agent of rice blast, another serious disease in rice. *Xa21*, described above, encodes a receptor kinase (Song et al., 1995); *Xa1* and *Pib* belong to the NBS–LRR class of *R* genes (Yoshimura et al., 1998; Wang et al., 1999). Various methods, including map-based cloning and transposon tagging, are currently being employed to isolate additional rice *R* genes (Ronald, 1997). At the same time, PCR strategies that amplify conserved motifs have led to the identification of several rice NBS–LRR *R* gene homologues (Leister et al., 1998). Cloning additional rice *R* genes will not only further our understanding of disease resistance mechanisms, especially in monocot systems, but may also have agronomic value since gene-for-gene *R* genes are the major tools of plant breeders to construct resistant varieties.

Because rice cultivar IR64 is an economically important high-yielding variety that is planted all over Asia, we wanted to characterize its array of *R* gene homologues with the assumption that some will correspond to disease resistance loci that have been identified previously. In this paper, we report the isolation and characterization of full-length homologues of *R* genes in rice cultivar IR64 that were identified using the *Arabidopsis* NBS–LRR *RPS2* gene (Bent et al., 1994; Mindrinos et al., 1994) as a heterologous hybridization probe.

2. Materials and methods

2.1. DNA and RNA methods

Standard protocols were used for restriction enzyme digestions, agarose gel electrophoresis, subcloning, and RNA and DNA gel blot analysis (Ausubel et al., 1999). Low stringency DNA blot analysis using an *RPS2* DNA probe and rice genomic DNA was carried out overnight at 42°C in 1% SDS, 2× SSC, and 10% dextran sulfate. Washing was performed as follows: 10 min in 2× SSC at room temperature, twice for 30 min each in 1% SDS, 6× SSC at 50°C. DNA sequencing was carried out using the T7 sequenase quick-denature plasmid sequencing kit (Amersham Life Science). PCR products to be sequenced were first cloned using the TA cloning kit (Invitrogen). DNA fragments were eluted from agarose gel slices using a Geneclean kit (Bio 101, Inc.).

Large amounts of rice genomic DNA for gel blot analysis and library construction were isolated using the method of Dellaporta et al. (1983), while crude DNA samples suitable for PCR were prepared using the method of Huang et al. (1997).

To obtain *NRH1*, *NRH2*, and *NRH3* sequence, relevant fragments were isolated from BAC clones 42A1, 37B13, and 44C24 (Yang et al., 1997), respectively, subcloned into Bluescript KS (+) (Stratagene), and sequenced using standard procedures (Ausubel et al., 1999). DNA sequences from rice cultivars other than IR64 and IR72 were obtained from PCR products. A 215 bp region around the premature stop codon in *NRH3* in cultivar IR64 was amplified from cultivars IR24, IRBB3, and Zenith using primers R3F (5′-GACCATTCACAGGCAGCTAC-3′) and R3R (5′-GATGCTATCCTCATCTCCCC-3′), and then sequenced. An 831 bp region in the open reading frame (ORF) adjacent to *NRH2* was amplified from IRBB3 using as primers the synthetic oligonucleotides 5′-AAATTCTCGAGCTTGTCAGC-3′ and 5′-GATCTCTGACAATCTTTGGG-3′, and sequenced.

Total RNA was prepared from leaves of 8 week-old IR64 plants using the RNAgents total RNA isolation system (Promega). Total RNA was prepared from TP-309 transgenic plants using the total RNA isolation system (Life Technologies). Poly(A)$^+$ RNA was prepared using the PolyATtract mRNA isolation system (Promega), and cDNA was synthesized by the Universal Riboclone cDNA synthesis system (Promega). To confirm the putative intron in *NRH1* and *NRH2*, sequences flanking the intron were used as primers in RT-PCR reactions using the cDNA preparation as template. PCR conditions were as follows: initial denaturation at 94°C for 3 min; amplification using 40 cycles of 94°C for 30 s, 55°C for 30 s, and 72°C for 60 s; and final extension at 72°C for 10 min. Primers for *NRH1* were 5′-GTCTCGATGCCGCCATTAAC-3′ and 5′-ATCCCATGTGCATATATCCC-3′. Primers for *NRH2* were 5′-AACTTGCTCGATGTGTGCGG-3′ and 5′-ATTACCATCCGGCAAAGCGC-3′. Only products of the expected size (for *NRH1*, 0.3 kb vs. 0.5 kb for corresponding fragment in hnRNA and genomic DNA, and for *NRH2*, 0.3 kb vs. 1.0 kb for corresponding fragment in hnRNA and genomic DNA) were obtained, and these were sequenced as described above. As further confirm-

ation of *NRH1* and *NRH2* expression, 256 bp from the 5′ end of the *NRH1* cDNA and 320 bp from the 3′ end of the *NRH2* cDNA were amplified using the above RT-PCR conditions (except that the annealing temperature was 50°C instead of 55°C), and sequenced. Primers for *NRH1* were 5′-AAGTACACCTTGGATATCTG-3′ and 5′-TCCAGCACCTGACTACCATG-3′. Primers for *NRH2* were 5′-CAGGACGTCCTAGCGTCAAC-3′ and 5′-GGTAAGCGCTGCCACAATAC-3′.

pNRH1 was constructed by inserting a 6 kb *Sal*I fragment from BAC clone 42A1 containing the *NRH1* gene into the *Sal*I site of pCAMBIA1301 (CAMBIA). pNRH2 was constructed by ligating a 9 kb *Kpn*I/*Xba*I fragment from BAC clone 37B13 containing the *NRH2* gene with *Kpn*I/*Xba*I-cut pCAMBIA1301.

2.2. Isolation of NRH clones

A subgenomic library was constructed from 2–5 kb *Eco*RI fragments (in which fraction some of the *RPS2*-hybridizing bands were observed) of IR72 DNA in λgt11, according to the instructions of the manufacturer (Stratagene). This library, which had a size of 3.6×10^4 plaques, was amplified and then screened using *RPS2* as probe under low stringency hybridization conditions (5× Denhardt's, 5× SSC, 0.2% SDS at 42°C) overnight, then rinsed twice for 5 min each in 2× SSC at room temperature, and washed for 30 min with 2× SSC, 0.1% SDS starting at 42°C and then at progressively higher temperatures (55 and 65°C) until some plaques gave a hybridization signal above background. Isolation and purification of positive plaques as well as preparation of λ DNA were performed using standard protocols (Ausubel et al., 1999). *NRH* clones from IR64 were isolated from a BAC library of IR64 DNA (Yang et al., 1997) by using PCR to screen DNA pools that were prepared from the BAC clones as described by Xu et al. (1998). Oligonucleotides derived from the sequence of *NRH1* in IR72 (5′-ATGGGATATATGCACATGGG-3′ and 5′-ATATCCAAGTGTACGAAGCG-3′ corresponding to nucleotides 1152–1171 and 2309–2328, respectively, in Fig. 1) were used as primers to amplify a 1.2 kb fragment in the PCR reactions.

2.3. Analysis of DNA and protein sequences

DNA and protein sequences were analyzed using programs of the University of Wisconsin Genetics Computer Group (Devereux et al., 1984). Per cent identities and similarities between sequences were calculated using the BESTFIT program. Homology searches were performed using the BLAST program, and the GRAIL program assisted in the search for ORFs. For the alignment shown in Fig. 4, MEGALIGN in the LASERGENE software for Macintosh (DNASTAR, Inc.) was used.

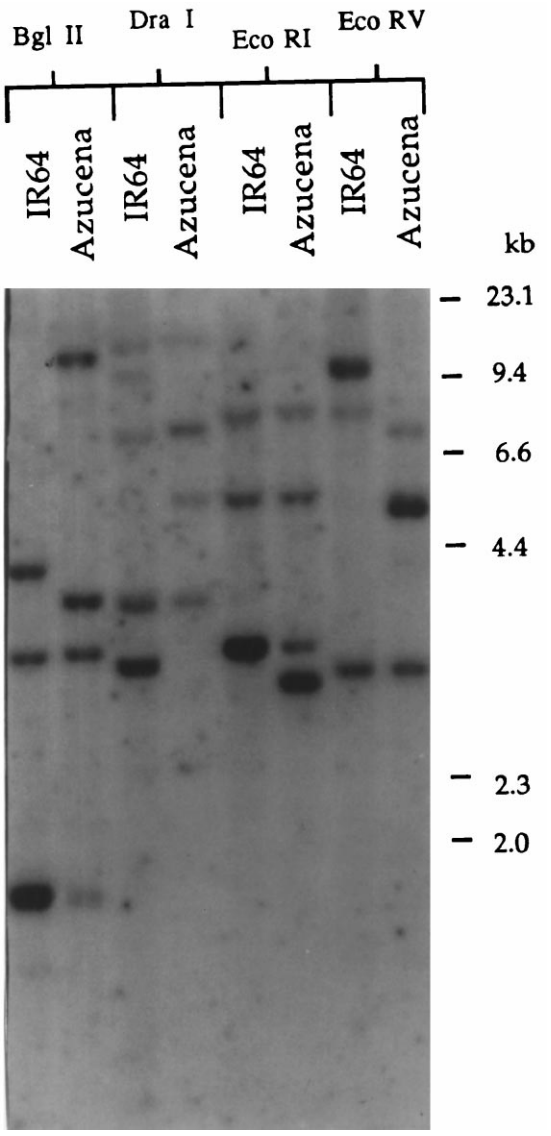

Fig. 1. DNA gel blot analysis of *NHR*. Genomic DNA (5 μg/lane) of the indicated rice cultivars was digested with the indicated restriction enzymes, fractionated by electrophoresis on a 0.8% agarose gel, and transferred to Hybond-N nylon membrane (Amersham Life Science, Inc.). The blotted filter was probed with the 3.5 kb *Eco*RI fragment isolated from IR72 containing part of *NRH1*, under high stringency conditions. Hybridization was performed in 5× SSC, 0.5% SDS, 5× Denhardt's, and 20 μg/ml sheared salmon sperm DNA, at 65°C overnight. Washing was done as follows: twice for 5 min each in 2× SSC, 0.1% SDS at 65°C, once for 15 min in 1× SSX, 0.1% SDS at 65°C, and finally twice for 10 min each in 0.1× SSC, 0.1% SDS at 65°C.

2.4. Genetic mapping

NRH1 and *NRH3* revealed *Pst*I and *Bgl*II polymorphisms, respectively, between rice cultivars IR64 and Azucena. Filters containing *Pst*I- and *Bgl*II-digested DNAs of the core set of 60 lines of a doubled haploid population derived from IR64/Azucena (Huang et al., 1997) were hybridized with the 3.5 kb *Eco*RI fragment isolated from IR72 containing part of *NRH1*.

Segregating bands were scored as either 1 (IR64) or 3 (Azucena). The data set was combined with the RFLP data set (Huang et al., 1997), and linkage analysis was carried out using MAPMAKER (Lander et al., 1987) (version 2.0) on a Macintosh computer. Map units (cM) were derived using the Kosambi function (Kosambi, 1944).

2.5. Rice transformation

pNRH1 and pNRH2, plasmid constructs containing the *NRH1* and *NRH2* genes, respectively, were introduced into rice variety TP-309 following a modification of the method of Hiei et al. (1994). Seeds were dehusked and sterilized in 70% ethanol for 1 min and then in 20% sodium hypochlorite for 1 h, washed several times with sterile water, and cultured on callus induction medium (MS basal medium supplemented with 2.0 mg/l 2,4-D, 500 mg/l casamino acids, 500 mg/l proline, 30 g/l sucrose, and 2.5 g/l phytagel, pH 5.8). 4 week-old calli were used for transformation. They were removed from seeds and divided into small pieces (1–2 mm in diameter) and subcultured for 4–5 days on callus induction medium. Infection of calli was carried out using *Agrobacterium tumefaciens* strain EHA 105 containing pCAMBIA1301-based constructs that contain genes for hygromycin resistance and GUS. EHA105 was grown for 2 days at 28°C on AB medium supplemented with 50 mg/l hygromycin and 50 mg/l kanamycin. The bacteria were collected, resuspended in AAM medium (Toriyama and Hinata, 1985), supplemented with acetosyringone (100 μM), and allowed to grow for 1 h at 28°C. The calli were immersed in the bacterial suspension, swirled, and incubated for 30 min. They were then transferred to MS-AS medium (MS basal, 500 mg/l casamino acids, 500 mg/l proline, 2 mg/l 2,4-D, 30 g/l sucrose, 10 g/l glucose, 100 μM acetosyringone, 2.5 g/l phytagel, pH 5.2) and co-cultivated for 3 days in the dark at 26°C. The calli were washed with 250 mg/l carbenecillin, incubated on MS-CH selection medium (callus induction medium with 50 mg/l hygromycin and 250 mg/l carbenecillin, pH 5.8) in the dark at 26°C, and transferred to new plates every 2 weeks. Resistant calli which appeared after 4–6 weeks were transferred to regeneration medium (MS basal, 2.0 mg/l kinetin, 30 g/l sucrose, 50 mg/l hygromycin, 250 mg/l carbenecillin, 4.0 g/l phytagel, pH 5.8) and incubated under 16 h light/8 h dark cycle. Plants regenerated within 4 weeks. The putative transgenics were moved to magenta boxes for shoot elongation (MS basal, 30 g/l sucrose, 2.5 g/l phytagel, pH 5.8). After 2 weeks, the plants were transplanted into pots and grown in the greenhouse. DNA and RNA gel blot analyses, GUS assays, and PCR were performed to determine presence and expression of the transgene.

2.6. GUS assay

GUS expression in rice calli and plants was assayed as described by Jefferson et al. (1987) with X-gluc (5-bromo-4-chloro-3-indolyl glucuronide) as substrate. Calli were incubated at 37°C overnight with X-gluc solution in phosphate buffer.

2.7. Bacterial leaf blight infection tests

Resistance to bacterial leaf blight was determined by inoculation of fully-expanded leaves of 6 week-old rice plants with various *Xoo* strains using the leaf clipping method (Kauffman et al., 1973). 12 days after inoculation, plants were scored by measuring the lesion length.

3. Results

3.1. Isolation of initial NRH clones

A preliminary DNA gel blot analysis was performed to determine whether the *Arabidopsis* disease resistance gene *RPS2* cross-hybridized with rice genomic DNA. A blot of rice cultivar IR72 genomic DNA digested with various restriction enzymes was hybridized with a full-length *RPS2* cDNA probe under low stringency conditions as described in Section 2.2. Autoradiography revealed a number of distinct bands against a relatively uniform but smeared background (data not shown). Based on this analysis, an IR72 subgenomic library containing 2–5 kb *Eco*RI fragments was screened to identify clones that cross-hybridized with the *Arabidopsis RPS2* gene. Three different positive clones were eventually isolated, having *Eco*RI fragment insert sizes of 2.8, 3.5, and 4.5 kb. The 3.5 kb insert, which gave the strongest signal when hybridized to *RPS2*, was subcloned into Bluescript KS (+) (Stratagene), sequenced, and found to contain a putative ORF which BLAST analysis showed had significant homology with *Prf* [$P(N) = 1.7e-31$] and *RPM1* [$P(N) = 1.2e-29$], members of the NBS–LRR class of *R* genes (Grant et al., 1995; Salmeron et al., 1996), and the most homology with *b8* [$P(N) = 3.8e-105$], a barley gene related to the NBS–LRR type of *R* genes but whose function has not yet been elucidated (Leister et al., 1998). The rice NBS–LRR homologue was given the name *NRH1* (for NBS–LRR resistance gene homologue, number 1).

3.2. Isolation of NRH clones from IR64

As shown in Fig. 1, gel blot analysis of IR64 DNA using the *NRH1* clone isolated from IR72 as probe under high stringency conditions revealed three to five hybridizing bands, suggestive of a gene family with at least three closely related members. We screened a BAC

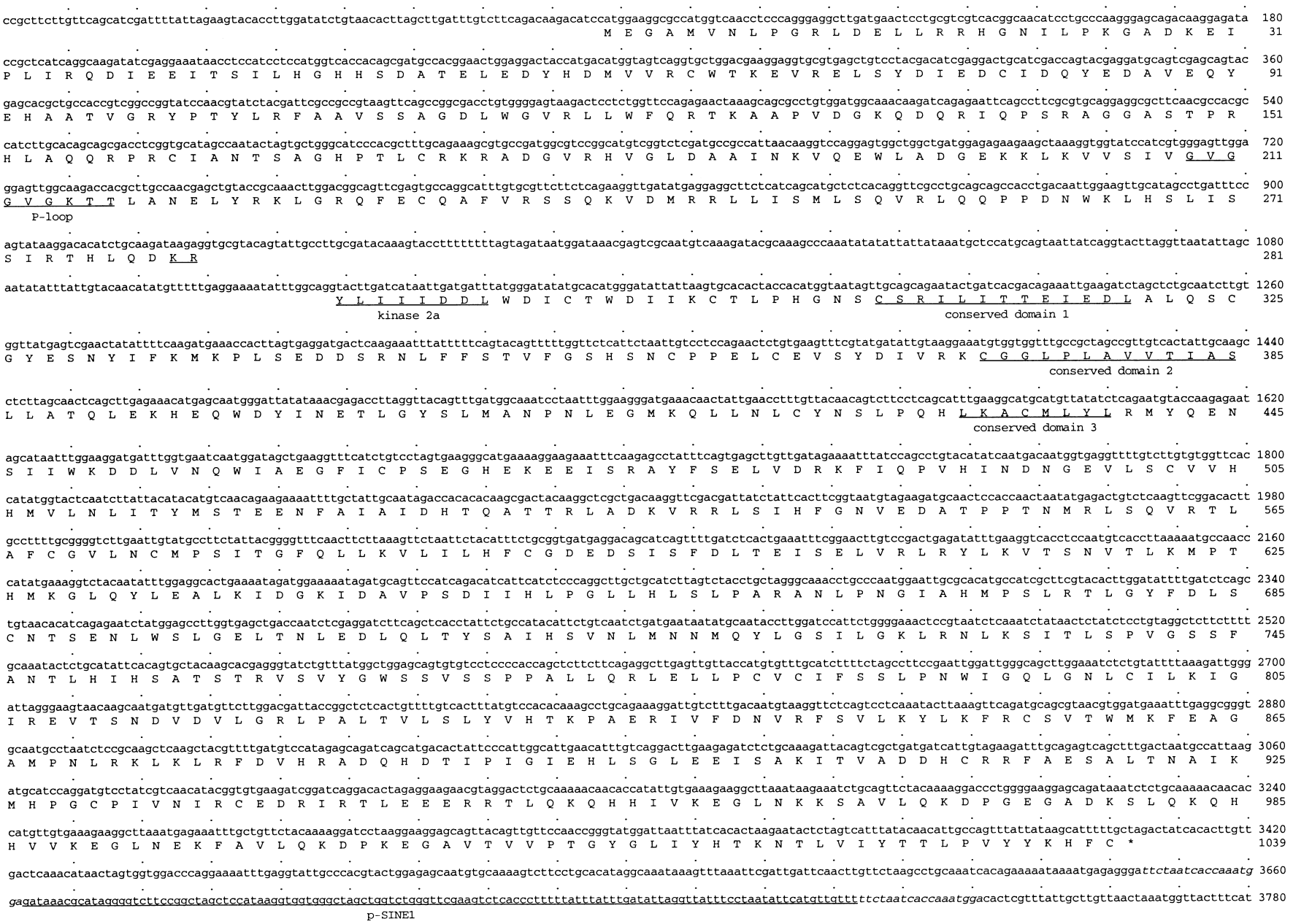

Fig. 2. *NRH1* sequence analysis. Conserved regions among NBS–LRR R proteins (P-loop, kinase 2a motif, and conserved domains 1,2, and 3) as well as a *p-SINE1* element are underlined. The 18 bp repeats flanking the *p-SINE1* element are italicized. Dots appear above every tenth nucleotide to facilitate location of a particular position. The LRR region is featured in Fig. 3.

library of IR64 DNA (Yang et al., 1997) for *NRH* member genes via PCR of BAC DNA pools, using primers derived from the IR72 sequence of *NRH1* (see Section 2.2 for details). Six positive BAC clones were identified, which on the basis of restriction enzyme mapping analysis (data not shown) were found to correspond to three of the hybridizing bands observed in the DNA blot analysis in Fig. 1. Four of the six BAC clones, 20J19, 24I9, 37L13, and 42A1, displayed similar *Bgl*II digestion patterns and contained a 1.6 kb *Bgl*II fragment which hybridized with *NRH1*, indicating that they represented the same region in the genome that corresponded to the 1.6 kb *Bgl*II hybridizing band shown in Fig. 1. Sequencing of the 1.6 kb fragment revealed 100% identity with the sequence of the *NRH1* gene previously isolated from IR72, indicating that the *NRH1* locus from IR64 had been cloned.

The two other BAC clones contained two different *NRH1*-related loci. BAC 37B13 and BAC 44C24 contained 3.5 and 4.3 kb *Bgl*II fragments, respectively, which hybridized to *NRH1*, corresponding to the similarly-sized *Bgl*II fragments observed in the DNA blot analysis in Fig. 1. We assigned the name *NRH2* to the *NRH* locus in 37B13, and the name *NRH3* to the *NRH* locus in 44C24. Sequencing the end of the insert in BAC 44C24 adjacent to the Sp6 promoter sequence in the vector pBeloBAC11 revealed that BAC 44C24 (which contains *NRH3*) also contains a portion of *NRH2* (no more than the first 400 base pairs). Because the size of the insert in BAC 44C24 is approximately 48 kb, *NRH2* and *NRH3* are at most 48 kb apart.

We screened a cDNA library prepared from leaf mRNA of greenhouse-grown, 8 week-old IR64 rice plants for *NRH* clones but were not able to isolate any hybridizing clones, suggesting that *NRH* mRNA levels might be low under these conditions. Indeed, no *NRH* messages were detected using RNA blot analysis of the same RNA fraction, even after prolonged exposure. RT-PCR analysis, however, yielded DNA bands corresponding to *NRH1* and *NRH2* (see Section 2.1 for details), confirming the likelihood that at least two of these genes were expressed, albeit at low levels (data not shown).

3.3. DNA sequence of IR64 NRH1, NRH2, and NRH3

We obtained 5.6 kb of cultivar IR64 genomic sequence around the *NRH1* locus, 10.2 kb of *NRH2* sequence, and 4.6 kb of *NRH3* sequence. The genomic *NRH1* sequence is shown in Fig. 2. The ORF encodes a putative polypeptide of 1039 amino acids with a molecular weight of 117 kDa. It shows very high homology (67% identity and 71% similarity at the amino acid level) with *b8*, a barley gene related to the NBS–LRR type of *R* genes but whose function has not yet been elucidated (Leister et al., 1998). Among *R* gene products whose biological function has been demonstrated, NRH1 shows the most homology with *Arabidopsis* RPM1 (28% identity, 40% similarity), tomato Prf (28% identity, 41% similarity), and rice Pib (30% identity, 41% similarity), members of the NBS–LRR class. As shown in Fig. 3, NRH1 has the five conserved domains shared by the NBS–LRR class, as well as an LRR region. Immediately downstream of the LRR region near the C-terminus are two copies of a stretch of 30

```
Consensus   LXXLXXLXXL/IXa

     559    LSQVRTLAFCGV  LNCMPSITGFQ
     582    LLKVLILH FCG  DEDSISFDLTEISE
     607    LVRLRYLK VTS  NVTLKMPTHMKG
     630    LQYLEALK IDG  KIDAVPSDIIH
     652    LPGLLHLS LPA  RANLPNGIAHMPS
     676    LRTLGYFD LSC  NTSENLWSLGELTN
     701    LEDLQ LT YSA  IHSVNLMNNMQYLGSI
     727    LGKLRNLKSITL  SPVGSSFANTLHIHSATSTRVSVYGWSSVSSPPAL
     774    LQRLELLPCVCI  FSSLPNWIGQ
     796    LGNLCILK IGI  REVTSNDVDV
     817    LGRLPALTVLSL  YVHTKPAERIVFDNVRFSV
     848    LKYLKFRCSVTW  MKFEAGAMPN
     870    LRKLK LRFDVH  RADQHDTIPIG
     892    IEHLSGLEEISA  KITVADDHCRRFAESALTNAIKMHPGCP
     932    IVNIRCEDRIRTL EEERR

     950    TLQKQHHIVKEGLNKKSAVLQKDPGEGADK
              |||||| |||||| | ||||||| |||
     980    SLQKQHHVVKEGLNEKFAVLQKDPKEGAVT
```

Fig. 3. LRR and repeat regions of NRH1. The amino acid sequence of the NRH1 LRR (amino acids 559–949) and the repeat region (amino acids 950–1009) are shown. The top line indicates the consensus sequence for the NRH1 LRR. An 'X' represents any amino acid, while an 'a' stands for an aliphatic residue. A vertical bar represents identity between the two sequences at that particular position.

amino acids, arranged in tandem, with 77% amino acid identity between the two copies.

The *NRH1* coding sequence is interrupted by a single intron, confirmed by RT-PCR, that is 198 nucleotides long (positions 930–1127 in Fig. 2). The intron follows the GT–AG rule, beginning and ending with the dinucleotides GT and AG, respectively. A short interspersed element (SINE) retroposon is present 198 bp downstream of the stop codon (positions 3603–3723 in Fig. 2). It shows the most homology with *r3* (87% identity) of the p-SINE1 family (Mochizuki et al., 1992). It is 121 bp in length and flanked by 18 bp direct repeats.

NRH2 encodes a polypeptide of 922 amino acids, slightly shorter than NRH1, with a molecular weight of 104 kDa. NRH2 is very similar to NRH1, as shown in the alignment in Fig. 4, with 76.7% identity, and 80% similarity. Like NRH1, NRH2 possesses domains characteristic of NBS–LRR R proteins. The stretch of 30 amino acids that is duplicated in NRH1 is present but not repeated in NRH2. *NRH2* also has an intron, confirmed by RT-PCR, at the same location relative to the amino acid sequence (KR↓YLII) as NRH1, a further indication of a close evolutionary relationship between *NRH1* and *NRH2*. The intron follows the GT–AG rule, and is 633 nucleotides in length, considerably bigger than the intron in *NRH1*.

It is interesting to note that in addition to *NRH2*, GRAIL and BLAST both identify a non-overlapping ORF that runs in the opposite direction of *NRH2* only 1.6 kb upstream of *NRH2* that also encodes an NBS–LRR protein. Although this second ORF appears to be prematurely truncated by a stop codon because of a frameshift, a polypeptide of 815 amino acids can be defined by a continuous chimera of different reading frames that circumvent the premature stop codons. Although it shows significant homology with the NBS–LRR class of R proteins, the resulting chimera is less homologous to NRH2 (30% identity, 39% similarity) than NRH1 is to NRH2 (77% identity, 80% similarity).

Upon sequencing, the 4.3 kb *Bgl*II fragment subclone from BAC 44C24 was found not to contain the entire *NRH3* ORF. The coding sequence for the kinase 2a domain starts at position 2734 in the 4.3 kb DNA sequence, upstream of which we were not able to find any additional potential exons. This suggested that the coding sequence corresponding to the amino terminus of the encoded polypeptide is either separated from the rest of the coding sequence by a large intron (>2.7 kb) or is absent altogether in this particular *NRH3* allele. We have also detected a premature stop codon generated by a frameshift only 825 bp downstream of the kinase 2a motif coding sequence. Hence the *NRH3* gene in IR64 could at best only encode a truncated and most likely a non-functional polypeptide. This being the case, we did not see any compelling reason to obtain further sequence of the locus.

3.4. Mapping of NRH1, NRH2, and NRH3

Since the *NRH* loci revealed RFLPs between IR64 and Azucena (Fig. 1), we were able to use the doubled haploid population derived from IR64/Azucena to determine the chromosomal locations of *NRH1*, *NRH2*, and *NRH3* as described in Section 2.4. *NRH1* was mapped 5.2 cM north of RZ70 and 15.4 cM south of RZ67 on chromosome 5, while *NRH3* was mapped 5.0 cM north of NpB186 on chromosome 11, as shown in Fig. 5. We were not able to identify an RFLP associated with *NRH2*, but since *NRH3* and *NRH2* are physically linked in a single BAC clone, it is reasonable to assume that *NRH2* maps to the same site as *NRH3*. It is interesting to note that these *NRH* loci map to regions where loci important for resistance to *Xoo* have been mapped previously. When the same doubled haploid population we used for mapping the *NRH* loci was analyzed by Huang and colleagues for horizontal resistance to *Xoo*, they were able to identify a QTL for resistance on chromosome 5 linked to marker RZ67 (Huang, unpublished data). Meanwhile, it is well established that IR64 has a functional *Xa4* gene, a major resistance gene to *Xoo*, which maps to the same region on chromosome 11 (Yoshimura et al., 1995) as *NRH2* and *NRH3*. Yoshimura et al. (1995) mapped the *Xa4* locus 1.7 cM north of NpB186, while in this report we have calculated *NRH3* to be 5.0 cM north of NpB186. In addition, a small F2 population (38 individuals) from a cross between near isogenic lines IR24(*xa4*/*xa4*) and IRBB4(*Xa4*/*Xa4*) showed complete co-segregation of the *Xa4* resistance phenotype and an RFLP detected by *NRH* (data not shown), suggesting tight linkage (approximately 1.3 cM) of *Xa4* with *NRH2*/*3*. Further, PCR analysis of IRBB4 cDNA, but not of IR24 cDNA, yielded the correct product using *NRH2*-specific primers (data not shown), indicating that *NRH2* is contained in the fragment containing *Xa4* that was introgressed into IR24 in generating IRBB4.

3.5. Bacterial blight infection tests of NRH1 and NRH2 transgenic plants

NRH1 and *NRH2* transgenes were introduced into rice variety TP-309 in an effort to elucidate their biological function. TP-309 was chosen because it is easy to transform and is susceptible to most *Xoo* strains. Expression of the *NRH1* or *NRH2* transgene was determined by PCR, as shown in Fig. 6. Transformants (T0 generation) that were shown to contain and express the *NRH1* or *NRH2* transgene were allowed to self and set seed. Plants grown from the seeds (T1 generation) were then tested for resistance to race 1 (Pxo 61 and Pxo 35) and race 5 (Pxo 112 and 105) *Xoo* strains. At the same time, segregation of the transgene was determined. None of the lines that express either the *NRH1* or *NRH2*

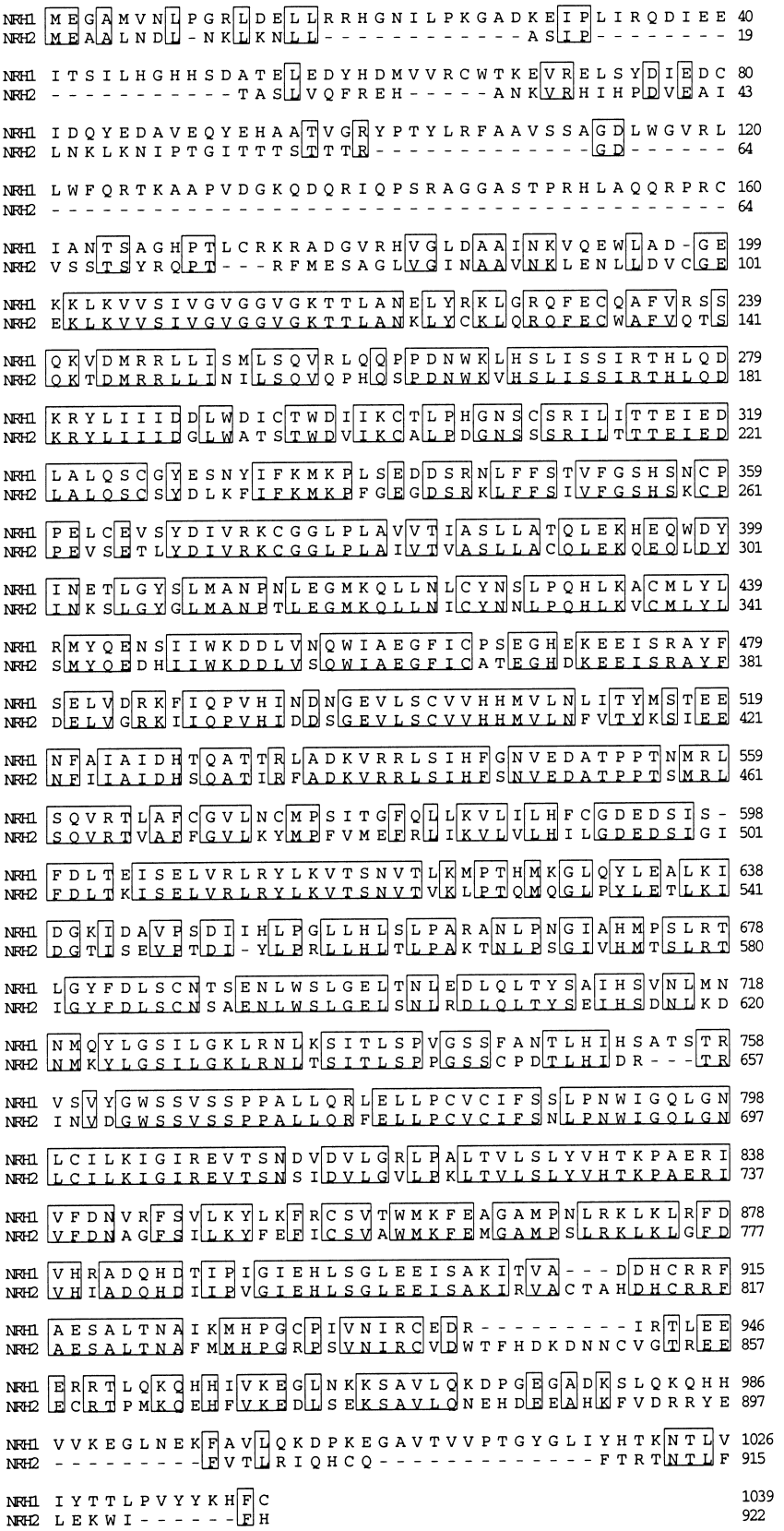

Fig. 4. Alignment of NRH1 and NRH2. The alignment was obtained by the Clustal method using MEGALIGN in the LASERGENE software for Macintosh (DNASTAR, Inc.). Positions with identical amino acids are boxed.

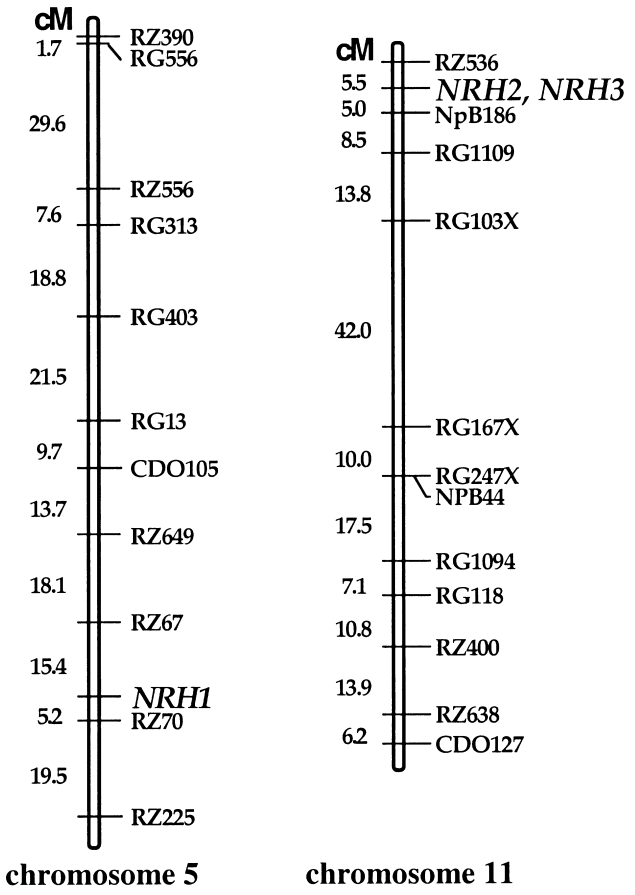

Fig. 5. Chromosomal locations of *NRH1*, *NRH2*, and *NRH3*. Genetic maps of chromosomes 5 and 11 are shown to indicate map locations of *NRH1*, *NRH2*, and *NRH3*. *NRH1* and *NRH3* were mapped based on linkage of RFLPs they detected with RFLP markers previously mapped in the same doubled haploid population derived from an IR64/Azucena cross (Huang et al., 1997). *NRH2* is at most 48 kb away from *NRH3* in IR64, and is therefore assumed to be linked to *NRH3*.

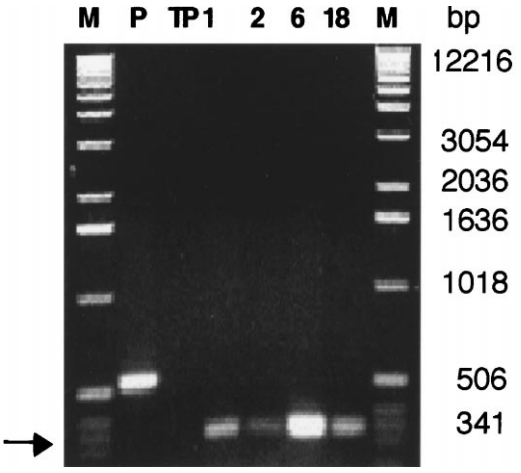

Fig. 6. Determination of *NRH1* transgene expression in primary transformants. Electrophoretic separation of products from PCR performed on cDNA prepared from the indicated transgenic lines (1, 2, 6, and 18) and the parental line TP-309 (TP) using primers that flank the *NRH1* intron (see Section 2.1) is shown. Column marked 'P' shows the PCR product when pNRH1 is used as template. In this case, the PCR product is larger compared with those from the transgenic lines because it includes the intron. Columns marked 'M' are DNA molecular size markers with fragment length in base pairs indicated on the right.

transgene showed resistance to the *Xoo* strains tested (data not shown), suggesting that NRH1 and NRH2 do not serve to confer resistance to these *Xoo* strains.

4. Discussion

We report here the identification of a family of *R* gene homologues that belongs to the NBS–LRR class of *R* genes, and the isolation of three members of the family from rice variety IR64. Interestingly, the isolated genes show slightly more homology with *RPM1* than *RPS2*, the probe initially used in screening the subgenomic library. This is somewhat surprising because *RPM1* and *RPS2* do not cross-hybridize, even at low stringency (Dewdney and Ausubel, unpublished data), although both are members of the NBS–LRR family of *R* genes.

Previously studied NBS–LRR *R* genes have also been shown to be members of gene families, including the *I2C-1* gene of tomato that confers resistance to *Fusarium oxysporum* f sp. *lycopersici*, the causal agent of vascular wilt disease. *I2C-1* is a member of the complex *I2C* that has been localized in five genomic positions, two of which are clusters of several genes on chromosome 11 (Ori et al., 1997). The *SL8D* cluster on chromosome 11 contains at least four genes. *I2C* family members also have tandem repeats of an almost identical sequence close to the C-terminus (Ori et al., 1997), similar to those found downstream of NRH1. These repeats have not been observed in other NBS–LRR R proteins, and their biological significance is presently unknown.

Naturally-occurring *R* gene alleles that encode truncated polypeptides such as *NRH3* appear to be relatively common. For example, the structure of the tomato *I2C-4* gene is similar to *NRH3* and the truncated *R* gene adjacent to *NRH2*. All of these genes contain frameshifts that result in putative translation of truncated peptides but whose full-length product may be represented by a continuous chimera of different reading frames (Ori et al., 1997). Similarly, a 90 kb region on chromosome 4 of *Arabidopsis* contains eight homologues of the disease resistance gene *RPP5*, a member of the NBS–LRR family that confers resistance to the oomycete fungus *Peronospora parasitica* race Noco2 (EU Arabidopsis Genome Project, 1998). However, most of these homologues apparently do not encode functional polypeptides, for a variety of reasons: frameshift mutations that prematurely terminate the ORF, retrotransposon insertions in the coding region, or the lack of an in-frame ATG initiator codon. It is possible that in other rice varieties, the *R* genes that are defective in IR64 may

encode full-length polypeptides with novel disease resistance specificities. Indeed, in other rice varieties like IRBB3, we have detected single-base insertions that, among other things, abolish the frameshift in *NRH3*, and some of the frameshifts in the gene adjacent to *NRH2*. However, we have not obtained enough sequence to indicate that in IRBB3 these genes could encode polypeptides whose length would be comparable with other R proteins of the NBS–LRR type, such as RPM1 and RPS2 (Ilag, unpublished data).

The fact that *R* genes are under constant evolutionary pressure to recognize novel *avr* gene products suggests that *R* gene loci may exhibit relatively high rates of recombination and mutation to generate novel *R* specificities. From a practical plant breeding standpoint, *R* genes or their homologues may be a rich source of polymorphisms from which PCR or RFLP markers may be developed for a variety of basic (e.g. genetic mapping) and applied (e.g. marker-aided selection) research objectives. We have already seen that the *NRH* loci reveal numerous RFLPs between IR64 and Azucena. Furthermore, the frameshift that causes a premature stop codon in *NRH3* in IR64 is not present in IRBB3, which creates a *Tsp509*I polymorphism. Oligonucleotides R3F and R3R (see Section 2.1) flank this polymorphism, and we exploited this sequence difference as a cleaved amplified polymorphic sequence (CAPS) marker (Konieczny and Ausubel, 1993) (data not shown). Another polymorphism that we have detected results in PCR products of different sizes in IR64 and IRBB3. In a PCR reaction, oligonucleotides (5′-TGCCTAATGGGATCCAGACTGGCGCAGCAC-3′ and 5′-GCAGCTTTCATGACATCGGCGTCTATACCG-3′) flanking the region separating *NRH2* and the adjacent potential ORF amplify a 1.6 kb fragment from IR64 DNA, and a 2.2 kb fragment from IRBB3 DNA. Because these PCR markers map in a region on chromosome 11 where a large collection of resistance specificities has been localized, they may be of potential use in mapping these *R* genes, or for marker-aided selection of plants carrying these *R* genes. In addition, because *p-SINE* elements have been shown to be polymorphic (Mochizuki et al., 1992), there is the possibility that the *p-SINE* element downstream of *NRH1* in IR64 may not be present in other rice varieties, so that cultivar-specific PCR markers may also be developed to tag this particular location on chromosome 5.

Given the high degree of homology shared by NRH1 and NRH2 with the NBS–LRR type of R proteins, it is likely that they function as disease resistance proteins. Even if they do, it may be quite challenging to identify the pathogen(s) [and subsequently the *avr* product(s)] that NRH1 and NRH2 correspond to. One possible method for matching *R* genes and their corresponding pathogens is to identify *R* genes that co-segregate with resistance to a particular pathogen. Thus we have shown that *NRH1* maps to a region on chromosome 5, where a QTL for horizontal resistance to *Xoo* has been identified, and *NRH2* is closely linked to *Xa4*, a major resistance gene that confers resistance to *Xoo*. These results raise the possibility that *NRH1* and *NRH2* confer resistance to *Xoo*. *Xoo* resistance is governed by *R–avr* gene interactions, and it is likely that some of the genes that confer resistance to *Xoo* will be of the NBS–LRR variety. Indeed, *Xa1*, a major *Xoo* resistance gene that was cloned recently, belongs to the NBS–LRR class of *R* genes (Yoshimura et al., 1998).

In an attempt to identify the putative pathogens that correspond to *NRH1* and *NRH2*, transgenic rice plants were generated and studied with respect to their disease susceptibility/resistance to infection by *Xoo*. However, the transgenic lines did not exhibit resistance to the *Xoo* strains tested which are normally avirulent on *Xa4*-containing plants. NRH2 therefore does not appear to correspond to Xa4. One explanation for the failure of *NRH1* and *NRH2* to confer resistance to the *Xoo* races tested is that TP-309 lacks other components of the signal transduction mechanism leading to NRH1- or NRH2-mediated resistance, such that susceptibility is still observed even with *NRH1* or *NRH2* present. However, this seems unlikely as previous reports show that disease resistance may be conferred on a different cultivar (Song et al., 1995) or a different plant species altogether (Song et al., 1995; Whitham et al., 1996) with the introduction of a foreign *R* gene, indicating that the basic signal transduction pathway leading to disease resistance is conserved in plants. Another possibility is that the *NRH1* and *NRH2* transgenes are not expressed at sufficiently high levels in TP-309, although, as shown in Fig. 6, transgenic lines were screened for ones that exhibited *NRH1* and *NRH2* expression.

The simplest interpretation of our data is that *NRH1* and *NRH2* are located in *R* gene clusters and encode functional resistance gene products, but that these products do not confer resistance to the limited number of pathogen races tested as part of this study. It would therefore be of interest to determine whether other *R* genes are located in the vicinity of the *NRH* loci. The DNA sequence of the rice genome, which is currently being determined by the International Rice Genome Sequencing Project, will facilitate this type of analysis. The tip of chromosome 11, which contains *NRH2* and *NRH3*, is particularly interesting, since this area has a relatively high concentration of resistance genes, with at least nine major resistance genes and one QTL for resistance clustered in a 30 cM region (Ronald, 1997). However, since different members of the NBS–LRR class of resistance proteins confer resistance to a wide range of pathogens and pests, including viruses, bacteria, fungi, nematodes and insects, there is currently no simple method for identifying the putative pathogens/pests that correspond to *NRH1* and *NRH2*. Thus, at this point,

although the biological functions of *NRH1* and *NRH2* remain unresolved, the transgenic plants described in this report will provide a useful tool for correlating specific *R* genes with specific pathogens, for example, by determining their response to a more extensive panel of rice pathogens and pests.

Acknowledgement

We thank the late Delfin Lapis for IR72 seeds, and Daichang Yang, Jichen Xu, Jonaliza Lanceras, Gerard Magpantay, and Jessica Domingo for technical assistance. L.L.I. was supported by a NIH postdoctoral fellowship (NRSA no. 1F32GM16527). R.C.Y. was supported by a Rockefeller Foundation postdoctoral fellowship. Partial financial support from the Rockefeller Foundation to IRRI is greatly appreciated. This work was supported by NIH Grant GM48707 awarded to F.M.A.

References

Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., Struhl, K., 1999. Current Protocols in Molecular Biology. Greene Publishing Associates/Wiley Interscience, New York.

Bent, A.F., Kunkel, B.N., Dahlbeck, D., Brown, K.L., Schmidt, R., Giraudat, J., Leung, J., Staskawicz, B.J., 1994. *RPS2* of *Arabidopsis thaliana*: a leucine-rich repeat class of plant disease resistance genes. Science 265, 1856–1860.

Dellaporta, S.L., Wood, J., Hicks, J.B., 1983. A plant DNA minipreparation: version II. Plant Mol. Biol. Rep. 1, 19–23.

Devereux, J., Haeberli, P., Smithies, O., 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. 12, 387–395.

EU Arabidopsis Genome Project, 1998. Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of *Arabidopsis thaliana*. Nature 391, 485–488.

Grant, M.R., Godiard, L., Straube, E., Ashfield, T., Lewald, J., Sattler, A., Innes, R.W., Dangl, J.L., 1995. Structure of the *Arabidopsis RPM1* gene enabling dual specificity disease resistance. Science 269, 843–846.

Hiei, Y., Ohta, S., Komari, T., Kumashiro, T., 1994. Efficient transformation of rice (*Oryza sativa* L.) mediated by Agrobacterium and sequence analysis of the boundaries of the T-DNA. Plant J. 6, 271–282.

Huang, N., Parco, A., Mew, T., Magpantay, G., McCouch, S., Guiderdoni, E., Xu, J., Subudhi, P., Angeles, E.R., Khush, G.S., 1997. RFLP mapping of isozymes, RAPD and QTLs for grain shape, brown planthopper resistance in a doubled haploid rice population. Mol. Breed. 3, 105–113.

Jefferson, R.A., Kavanagh, T.A., Bevan, M.W., 1987. GUS fusions: beta-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. EMBO J. 6, 3901–3907.

Johal, G.S., Briggs, S.P., 1992. Reductase activity encoded by the *HM1* disease resistance gene in maize. Science 258, 985–987.

Jones, D.A., Thomas, C.M., Hammond-Kosack, K.E., Balint-Kurti, P.J., Jones, J.D.G., 1994. Isolation of the tomato *Cf-9* gene for resistance to *Cladosporium fulvum* by transposon tagging. Science 266, 789–793.

Kauffman, H.E., Reddy, A.P.K., Ksiek, S.P.V., Marca, S.D., 1973. An improved technique for evaluating resistance of rice varieties to *Xanthomonas oryzae*. Plant Dis. Rep. 57, 537–541.

Konieczny, A., Ausubel, F.M., 1993. A procedure for mapping *Arabidopsis* mutations using co-dominant ecotype-specific PCR-based markers. Plant J. 4, 403–410.

Kosambi, D.D., 1944. The estimation of map distance from recombination values. Ann. Eugen. 12, 172–175.

Lander, E.S., Green, P., Abrahamson, J., Barlow, A., Daly, M.J., Lincoln, S.E., Newburg, L., 1987. MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1, 174–181.

Leister, D., Kurth, J., Laurie, D.A., Yano, M., Sasaki, T., Devos, K., Graner, A., Schulze-Lefert, P., 1998. Rapid reorganization of resistance gene homologues in cereal genomes. Proc. Natl. Acad. Sci. USA 95, 370–375.

Martin, G.B., Brommonschenkel, S.H., Chunwongse, J., Frary, A., Ganal, M.W., Spivey, R., Wu, T., Earle, E.D., Tanksley, S.D., 1993. Map-based cloning of a protein kinase gene conferring disease resistance in tomato. Science 262, 1432–1436.

Mindrinos, M., Katagiri, F., Yu, G.-L., Ausubel, F.M., 1994. The *A. thaliana* disease resistance gene *RPS2* encodes a protein containing a nucleotide-binding site and leucine-rich repeats. Cell 78, 1089–1099.

Mochizuki, K., Umeda, M., Ohtsubo, H., Ohtsubo, E., 1992. Characterization of a plant SINE, p-SINE*1*, in rice genomes. Jpn. J. Genet. 57, 155–166.

Ori, N., Eshed, Y., Paran, I., Presting, G., Aviv, D., Tanksley, S., Zamir, D., Fluhr, R., 1997. The *I2C* family from the wilt disease resistance locus *I2* belongs to the nucleotide-binding, leucine-rich repeat superfamily of plant resistance genes. Plant Cell 9, 521–532.

Ronald, P.C., 1997. The molecular basis of disease resistance in rice. Plant Mol. Biol. 35, 179–186.

Salmeron, J.M., Oldroyd, G.E.D., Rommens, C.M.T., Scofield, S.R., Kim, H.-S., Lavelle, D.T., Dahlbeck, D., Staskawicz, B.J., 1996. Tomato *Prf* is a member of the leucine-rich repeat class of plant disease resistance genes and lies embedded within the *Pto* kinase gene cluster. Cell 86, 123–133.

Song, W.-Y., Wang, G.-L., Chen, L.-L., Kim, H.-S., Pi, L.-Y., Holsten, T., Gardner, J., Wang, B., Zhai, W.-X., Zhu, L.-H., Fauquet, C., Ronald, P., 1995. A receptor kinase-like protein encoded by the rice disease resistance gene, *Xa21*. Science 270, 1804–1806.

Toriyama, K., Hinata, K., 1985. Cell suspension and protoplast culture in rice. Plant Sci. 41, 179–183.

Wang, Z.X., Yano, M., Yamanouchi, U., Iwamoto, M., Monna, L., Hayasaka, H., Katayose, Y., Sasaki, T., 1999. The Pib gene for rice blast resistance belongs to the nucleotide binding and leucine-rich repeat class of plant disease resistance genes. Plant J. 19, 55–64.

Whitham, S., McCormick, S., Baker, B., 1996. The *N* gene of tobacco confers resistance to tobacco mosaic virus in transgenic tomato. Proc. Natl. Acad. Sci. USA 93, 8776–8781.

Xu, J., Yang, D., Domingo, J., Ni, J., Huang, N., 1998. Screening for overlapping bacterial artificial chromosome clones by PCR analysis with an arbitrary primer. Proc. Natl. Acad. Sci. USA 95, 5661–5666.

Yang, D., Parco, A., Nandi, S., Subudhi, P., Zhu, Y., Wang, G., Huang, N., 1997. Construction of a bacterial artificial chromosome (BAC) library and identification of overlapping BAC clones with chromosome 4-specific RFLP markers in rice. Theor. Appl. Genet. 95, 1147–1154.

Yoshimura, S., Yoshimura, A., Iwata, N., McCouch, S.R., Abenes, M.L., Baraoidan, M.R., Mew, T.W., Nelson, R.J., 1995. Tagging and combining bacterial blight resistance genes in rice using RAPD and RFLP markers. Mol. Breed. 1, 375–387.

Yoshimura, S., Yamanouchi, U., Katayose, Y., Toki, S., Wang, Z.-X., Kono, I., Kurata, N., Yano, M., Iwata, N., Sasaki, T., 1998. Expression of *Xa1*, a bacterial blight resistance gene in rice, is induced by bacterial inoculation. Proc. Natl. Acad. Sci. USA 95, 1663–1668.